Exceptional service in the national interest

Sandia National Laboratories

Extreme Heterogeneity: Operating and Runtime Systems Issues (The Total Bear)

Ron Brightwell, R&D Manager

Scalable System Software Department

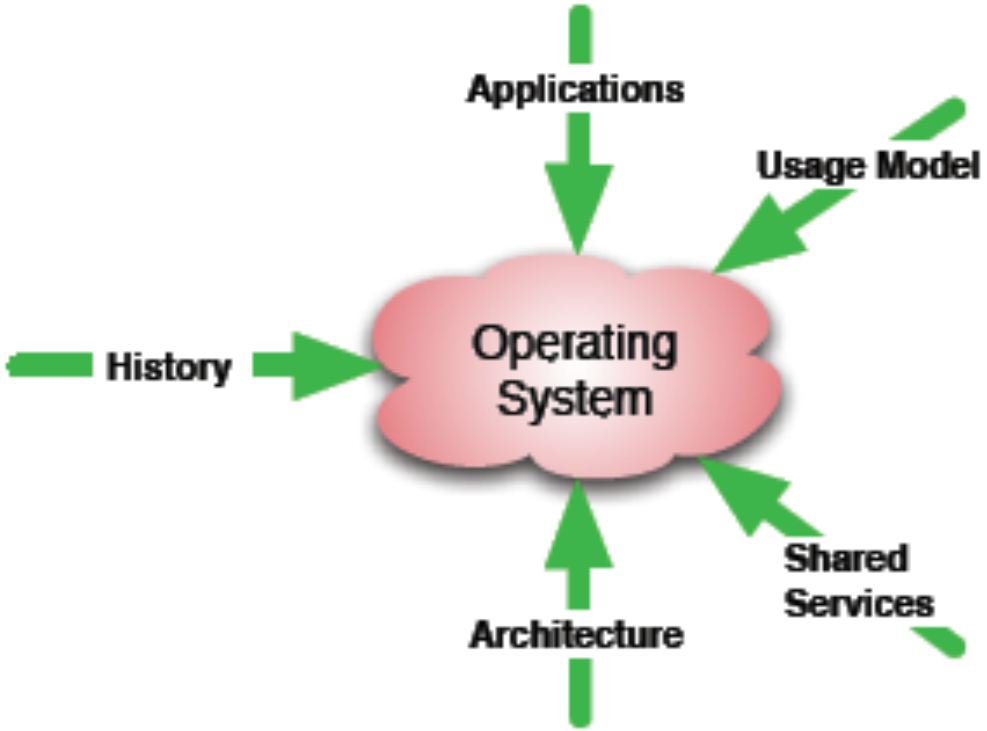
CCR Center for Computing Research

U.S. DEPARTMENT OF ENERGY | NNSA

Sandia National Laboratories is a multi-mission laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

Factors Influencing OS Design

Sandia National Laboratories



CCR Center for Computing Research

Architecture

Sandia National Laboratories

- System-on-Chip (SoC)
  - Hardware specialization
  - OS/R needs to be aware of custom hardware capabilities
    - Potentially large collection of hardware capabilities where only a few may be used at a time
  - A single node will not be a single cache-coherent physical address space (true today)
- Photonic interconnects
  - Load/store across a larger domain
  - More intelligent memory controllers
    - Perhaps programmable by OS/R or application
    - Converged with network interface
  - Nodes will look more like racks, racks will look more like systems
- Special-purpose systems will become more general
  - OS will have to be engineered to adapt more easily to new hardware
- Trust model will have to evolve
  - Security model for users and applications likely needs to change
- OS will become much more distributed

CCR Center for Computing Research

Applications

Sandia National Laboratories

- Increased complexity
  - Reduce complexity through componentization and composition
  - Decompose applications into tasks and services
  - OS/R will need to provide mechanisms for service discovery and composition
- Access to system services
  - Traps and blocking system calls are already insufficient
  - Convergence between OS and RTS
  - Expose hardware directly to application
- Tools are applications too
  - Tools typically depend more on system services
  - Less human interaction with tools
    - Consumer of diagnostic and debugging information may be the OS or RTS
- Rethink the connections between OS/R and programming environment
- Likely to be event-driven at some level

CCR Center for Computing Research

Usage Model

Sandia National Laboratories

- Need to move beyond batch-scheduled, space-shared, non-interactive jobs
  - Dedicated resources versus shared resources
  - More interactivity with users and application services
  - Need to develop a new cost charging model for facilities
- Implicit versus explicit allocation and management of resources
  - Already seeing limitations with explicitly allocating cores, nodes, memory (burst buffers) etc.
  - OS/R will likely need to determine resources implicitly and be elastic
  - Methods for handling resource failures
- Data-centric versus compute-centric view of system
  - Differentiating between HPC and Cloud/BigData approaches
- Support new methods of moving data on and off of the system

CCR Center for Computing Research

Shared Services

Sandia National Laboratories

- RAS System (Reliability/Availability/Serviceability)
  - System health monitoring
    - In-band and/or out-or-band
  - Global Information Bus
- External resources
  - External connectivity to network and storage
  - Streaming data from external instruments
  - New methods of data ingest/egest

CCR Center for Computing Research

History

Sandia National Laboratories

- Legacy programming interfaces
  - POSIX probably needs to go away for more than just I/O
  - Glibc may not be the RTS of the future
  - How to provide support for incremental adoption?
- Standard protocols
  - Which abstraction layers allow for evolution?
- May finally have to move away for Unix model
  - Convergence of memory and storage is a fundamental change for the OS
  - Everything is really not a file
- Need to balance between starting from scratch and supporting existing infrastructure

CCR Center for Computing Research